Optimality Analysis of A Class of Semi-infinite Programming Problems

Zhi Guo Feng · Fei Chen · Lin Chen · Ka Fai Cedric Yiu

Abstract In this paper, we consider a class of semi-infinite programming problems with a parameter. As the parameter increases, we prove that the optimal values decrease monotonically. Moreover, the limit of the sequence of optimal values exists as the parameter tends to infinity. In finding the limit, we decompose the original optimization problem into a series of subproblems. By calculating the maximum optimal values to the subproblems and applying a fixed-point theorem, we prove that the obtained maximum value is exactly the limit of the sequence of optimal values under certain conditions. As a result, the limit can be obtained efficiently by solving a series of simplified subproblems. Numerical examples are provided to verify the limit obtained by the proposed method.

Keywords Semi-infinite programming · Fixed-point theorem · Filter design · Beamformer design

Mathematics Subject Classification (2000) 90C34 · 47H10 · 26B10

Zhi Guo Feng,
Faculty of Mathematics and Computer Science, Guangdong Ocean University, Zhanjiang, Guangdong, P.R.C.
18281102@qq.com

Fei Chen
College of Mathematical Sciences, Chongqing Normal University, Chongqing, P.R.C.
835938789@qq.com

Lin Chen, Corresponding author
College of Mathematical Sciences, Chongqing Normal University, Chongqing, P.R.C. and College of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, Sichuan, P.R.C.
chinaallenchen@126.com

Ka Fai Cedric Yiu
Department of Applied Mathematics, Hong Kong Polytechnic University, Kowloon, Hong Kong, P.R.C.
macyiu@polyu.edu.hk

1 Introduction

Semi-infinite programming (SIP) problems [1–3] arise in many areas such as engineering design and game theory. There are various algorithms developed in the literature for solving SIP problems. They include discretization methods [2,3], gradient based methods [4], cutting plane methods [5,6] and exchange methods [7].

In certain applications such as [8,9] for filter or beamformer design, the objective function is related to some frequency response functions which in turn depend on the filter length. It is observed that when the filter length is sufficiently long, the optimal value could approach zero in some cases. However, this might not be generally true. It will be very useful to know and analyze the limit of the optimal value series as filter length increases. This will provide guidance to determine if a problem is posed properly and to set the filter length in order to achieve a desired performance. In view of this, a class of semi-infinite programming problems with one parameter L is considered in this paper. We aim at proposing an efficient method to find the limit of the optimal values for all possible L's.

The paper is organized as follows. In Section 2, a class of SIP problems is stated. In Section 3, we analyze the monotonicity of the optimal values and introduce a decomposition method to find the limit of the optimal values. A fixed-point theorem is applied to verify the limit. Numerical examples are provided in Section 4 and conclusions are summarized in Section 5.

2 Problem Formulation

We consider a class of SIP problems:

$$P(L) \qquad \begin{array}{l} \min\limits_{y,\boldsymbol{x}} y \\ g_i(\boldsymbol{H}(\boldsymbol{x},t)) - y \leq 0, \quad \forall t \in \Omega, i = 1,2,\cdots,m, \end{array}$$

where $H_j(\boldsymbol{x},t) = \sum\limits_{k=1}^{L} x_{jk}\varphi_k(t),\ j = 1,2,\cdots,N$, and

$$\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_N)^T, \quad \boldsymbol{x}_j = (x_{j1}, x_{j2}, \cdots, x_{jL}),$$

$$\boldsymbol{H}(\boldsymbol{x},t) = (H_1(\boldsymbol{x},t), \cdots, H_N(\boldsymbol{x},t))^T, \quad \boldsymbol{\varphi}(t) = (\varphi_1(t), \cdots, \varphi_L(t))^T.$$

The functions $\{\varphi_k(t), k = 1,2,\cdots\}$ form a basis, L is the number of selected functions in the basis and Ω is a given nonempty compact set in $\mathbb{R}$. The functions g_i are bounded from below and belong to $C^2(\mathbb{R}^N,\mathbb{R})$. Problem $P(L)$ is related to the parameter L. If the parameter L changes, the optimal value will also vary. In the following Section, the functional relationship between optimal value and the parameter L will be analyzed, and the best optimal value for all possible L's is sought.

3 Optimal Value Analysis

3.1 Optimal Value of the Problem

Let y^{*L} and $\boldsymbol{x}^{*L}$ denote the optimal value and the optimal solution of Problem $P(L)$ with parameter L, respectively. We have

$$y^{*L} = \max_{t\in\Omega} \max_{i\in\{1,2,\cdots,m\}} g_i(\boldsymbol{H}(\boldsymbol{x}^{*L}, t)).$$

First of all, the optimal value series satisfies monotonicity as shown below.

Theorem 3.1 *The optimal value series $\{y^{*L} : L = 1, 2, \cdots\}$ is monotonically decreasing and there exists a limit as L tends to infinity.*

Proof Define span(Δ_L) as the linear space spanned by the set Δ_L, where $\Delta_L := \{\varphi_1(t), \varphi_2(t), \cdots, \varphi_L(t)\}$. Then, we have span$(\Delta_L) \subset$ span(Δ_{L+1}) and $H_j(\boldsymbol{x}, t) \in$ span$(\Delta_L), \forall j \in \{1, 2, \cdots, N\}$.

If L increases, the feasible set of (y, H) increases monotonically. Hence, the optimal value series is monotonically decreasing. Moreover, for any i in $\{1, 2, \cdots, m\}$, g_i has a lower bound, therefore the optimal value series has a lower bound. Thus, the limit of the series $\{y^{*L} : L = 1, 2, \cdots\}$ exists. □

By Theorem 3.1, we denote the limit of $\{y^{*L} : L = 1, 2, \cdots\}$ by

$$y^* = \lim_{L\to\infty} y^{*L} = \inf_L y^{*L}.$$

Note that the limit y^* is the best value for all possible L. It can be used as the guidance in real application design. The target value of the design must be set to be poorer than y^*, and we can find a suitable L to satisfy the requirement. On the other hand, if the target value of the design is set to be better than y^*, we can not find any L to satisfy the requirement. In estimating the limit, a large L is always required. However, it also means that the implementation complexity and computational cost will also increase significantly. In the following, we propose an efficient method to tackle this problem.

3.2 Limit Value Analysis

In order to find the limit of Problem $P(L)$, the functional optimization problem is formulated as

$$\begin{aligned} &\min_{y, \overline{\boldsymbol{H}}(t)\in\Xi^N} \quad y \\ &g_i(\overline{\boldsymbol{H}}(t)) - y \le 0, \quad \forall t \in \Omega,\ i = 1, 2, \cdots, m, \end{aligned}$$

where $\overline{\boldsymbol{H}}(t) = (\overline{H}_1(t), \overline{H}_2(t), \cdots, \overline{H}_N(t))$, Ξ is the functional space given by

$$\Xi := \{Q(t) : \exists q^{(L)}(t) \in span(\Delta_L),\ s.t.\ Q(t) = \lim_{L\to\infty} q^{(L)}(t),\ \forall t \in \Omega\}.$$

However, this problem is highly complex. In order to lower the complexity, we propose the following decomposition method. For each $t \in \Omega$, $\overline{\boldsymbol{H}}(t) = (\overline{H}_1(t), \overline{H}_2(t), \cdots, \overline{H}_N(t))$ is treated as the decision vector and a subproblem can be formulated as

$$P(t) \qquad \begin{array}{l} \min\limits_{y, \overline{\boldsymbol{H}}(t)} \quad y \\ g_i(\overline{\boldsymbol{H}}(t)) - y \le 0, \quad i = 1, 2, \cdots, m. \end{array}$$

For each $t \in \Omega$, Problem $P(t)$ is a general optimization problem. Denote the optimal solution and the optimal value of Problem $P(t)$ by $\overline{\boldsymbol{H}}_t^*$ and y_t^*. Then, $\overline{\boldsymbol{H}}_t^*$ can be united as a function $\overline{\boldsymbol{H}}^*(t)$ with respect to t, and

$$y^{*t} = \max_{t \in \Omega} y_t^*. \tag{1}$$

In the following, we aim at finding the conditions such that $y^* = y^{*t}$. Consider the KKT optimality conditions of Problem $P(t)$ as

$$\begin{array}{l} \nabla_{\boldsymbol{w}} y + \sum\limits_{i=1}^{m} \lambda_i \nabla_{\boldsymbol{w}} \left(g_i(\overline{\boldsymbol{H}}(t)) - y \right) = 0, \\ \lambda_i \left(g_i(\overline{\boldsymbol{H}}(t)) - y \right) = 0, \\ g_i(\overline{\boldsymbol{H}}(t)) - y \le 0, \lambda_i \ge 0, \end{array} \tag{2}$$

where $\boldsymbol{w} := (y, \overline{\boldsymbol{H}}, \boldsymbol{\lambda}), \boldsymbol{\lambda} = (\lambda_1, \lambda_2, \cdots, \lambda_m)$. Let

$$\Gamma := \{\boldsymbol{w} = (y, \overline{\boldsymbol{H}}, \boldsymbol{\lambda}) : g_i(\overline{\boldsymbol{H}}) - y \le 0, \lambda_i \ge 0, \forall i = 1, 2, \cdots, m\},$$

then, (2) is equivalent to the following equations

$$\boldsymbol{F}(t, \boldsymbol{w}^*) := \begin{pmatrix} 1 - \sum\limits_{i=1}^{m} \lambda_i \\ \sum\limits_{i=1}^{m} \lambda_i \left(\nabla_{\overline{\boldsymbol{H}}} g_i(\overline{\boldsymbol{H}}, t) \right)^T \\ \lambda_1 (g_1(\overline{\boldsymbol{H}}, t) - y) \\ \vdots \\ \lambda_m (g_m(\overline{\boldsymbol{H}}, t) - y) \end{pmatrix} = \boldsymbol{0}, \quad \boldsymbol{w^*} \in \Gamma, \tag{3}$$

where $\boldsymbol{w}^* = (y_t^*, \overline{\boldsymbol{H}}_t^*, \boldsymbol{\lambda}_t^*)$ is the optimal solution of Problem $P(t)$. For simplicity, we denote $g_i = g_i(\overline{\boldsymbol{H}})$. Note that $\boldsymbol{F}(t, \boldsymbol{w})$ is a $(N + m + 1)$-dimensional vector function, its gradient can be computed directly as

$$\nabla_{\boldsymbol{w}} \boldsymbol{F} = \nabla_{\boldsymbol{w}} \boldsymbol{F}(t, \boldsymbol{w}) = \begin{pmatrix} 0 & \boldsymbol{0} & -1 & -1 & \cdots & -1 \\ \boldsymbol{0} & \sum\limits_{i=1}^{m} \lambda_i \nabla^2_{\overline{\boldsymbol{H}}} g_i & \left(\nabla_{\overline{\boldsymbol{H}}} g_1\right)^T & \left(\nabla_{\overline{\boldsymbol{H}}} g_2\right)^T & \cdots & \left(\nabla_{\overline{\boldsymbol{H}}} g_m\right)^T \\ -\lambda_1 & \lambda_1 \nabla_{\overline{\boldsymbol{H}}} g_1 & g_1 - y & 0 & \cdots & 0 \\ -\lambda_2 & \lambda_2 \nabla_{\overline{\boldsymbol{H}}} g_2 & 0 & g_2 - y & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ -\lambda_m & \lambda_m \nabla_{\overline{\boldsymbol{H}}} g_m & 0 & 0 & \cdots & g_m - y \end{pmatrix},$$

where $\boldsymbol{0}$ is a zero vector with suitable dimension, $\nabla_{\overline{\boldsymbol{H}}} g_i = (\frac{\partial g_i}{\partial \overline{H}_1}, \frac{\partial g_i}{\partial \overline{H}_2}, \cdots, \frac{\partial g_i}{\partial \overline{H}_N})$. For Problem $P(t)$, we assume that the linearly independent constraint qualification ($LICQ$) hold for any $t \in \Omega$, that is,

Assumption 1 *$\forall t \in \Omega$, the gradients $\left\{\nabla_{\boldsymbol{w}}\left(g_i(\overline{\boldsymbol{H}}_t^*) - y_t^*\right), i \in I\right\}$ are linearly independent, where $I := \{i : g_i(\overline{\boldsymbol{H}}_t^*) - y_t^* = 0\}$.*

Based on the discussion above, we obtain the following lemma and theorem.

Lemma 3.1 *If y_t^* is the optimal value of Problem $P(t)$ for any $t \in \Omega$, then, there is at least one active constraint.*

Proof Assume that all the constraints are not active, that is, $g_i(\overline{\boldsymbol{H}}_t^*) - y_t^* < 0$, $\forall i = 1, 2, \cdots, m$. Define a new number y_0^* by

$$y_0^* = \max_{i \in \{1,2,\cdots,m\}} g_i(\overline{\boldsymbol{H}}_t^*) = g_k(\overline{\boldsymbol{H}}_t^*),$$

where $k \in \{1, 2, \cdots, m\}$ is the index which maximize the function. Then,

$$y_t^* > \max_{i \in \{1,2,\cdots,m\}} g_i(\overline{\boldsymbol{H}}_t^*) = y_0^*.$$

Hence, we have

$$g_p(\overline{\boldsymbol{H}}_t^*) - y_0^* = g_p(\overline{\boldsymbol{H}}_t^*) - \max_{i \in \{1,\cdots,m\}} g_i(\overline{\boldsymbol{H}}_t^*) \le 0, \ \forall p \in \{1, 2, \cdots, m\},$$

$$g_k(\overline{\boldsymbol{H}}_t^*) - y_0^* = g_k(\overline{\boldsymbol{H}}_t^*) - \max_{i \in \{1,\cdots,m\}} g_i(\overline{\boldsymbol{H}}_t^*) = g_k(\overline{\boldsymbol{H}}_t^*) - g_k(\overline{\boldsymbol{H}}_t^*) = 0.$$

It means that $(y_0^*, \overline{\boldsymbol{H}}_t^*)$ is also a feasible solution of Problem $P(t)$. However, $y_0^* < y_t^*$, which contradicts the optimality of y_t^*. Hence, the assumption does not hold and there is at least one active constraint. This completes the proof. □

Theorem 3.2 *Suppose that Assumption 1 holds, and* $\det\left(\sum_{i=1}^{m} \lambda_i \nabla^2_{\overline{\boldsymbol{H}}} g_i\right) \neq 0$ *at $(t, \boldsymbol{w}^*) = (t, y_t^*, \overline{\boldsymbol{H}}_t^*, \boldsymbol{\lambda}_t^*)$, then* $\det(\nabla_{\boldsymbol{w}} \boldsymbol{F}(t, \boldsymbol{w}^*)) \neq 0$.

Proof (i) First, we simplify the determinant $\det(\nabla_{\boldsymbol{w}} \boldsymbol{F}(t, \boldsymbol{w}^*))$.

For $\lambda_i(g_i(\overline{\boldsymbol{H}}_t^*) - y_t^*) = 0, i \in \{1, 2, \cdots, m\}$, there are two cases:

$$1).\ g_i(\overline{\boldsymbol{H}}_t^*) - y_t^* < 0, \quad \lambda_i = 0; \quad 2).\ g_i(\overline{\boldsymbol{H}}_t^*) - y_t^* = 0, \quad \lambda_i > 0.$$

Without loss of generality, we suppose that the first r $(r \ge 1)$ constraints are active while the other $m - r$ constraints are inactive. Then we can rewrite $\nabla_{\boldsymbol{w}} \boldsymbol{F}$ as

$$\nabla_{\boldsymbol{w}} \boldsymbol{F} = \begin{pmatrix} 0 & \boldsymbol{0} & -1 & \cdots & -1 & -1 & \cdots & -1 \\ \boldsymbol{0} & \sum_{i=1}^{m} \lambda_i \nabla^2_{\overline{\boldsymbol{H}}} g_i & (\nabla_{\overline{\boldsymbol{H}}} g_1)^T & \cdots & (\nabla_{\overline{\boldsymbol{H}}} g_r)^T & (\nabla_{\overline{\boldsymbol{H}}} g_{r+1})^T & \cdots & (\nabla_{\overline{\boldsymbol{H}}} g_m)^T \\ -\lambda_1 & \lambda_1 \nabla_{\overline{\boldsymbol{H}}} g_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ -\lambda_r & \lambda_r \nabla_{\overline{\boldsymbol{H}}} g_r & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & g_{r+1} - y & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & g_m - y \end{pmatrix}.$$

The determinant $\det(\nabla_{\boldsymbol{w}}\boldsymbol{F}(t,\boldsymbol{w}^*))$ can be simplified as an upper triangular determinant $\begin{vmatrix} A & * \\ 0 & B \end{vmatrix}$, where

$$|A| = \begin{vmatrix} 0 & \mathbf{0} & -1 & \cdots & -1 \\ \mathbf{0} & \sum_{i=1}^{m} \lambda_i \nabla^2_{\overline{\boldsymbol{H}}} g_i & (\nabla_{\overline{\boldsymbol{H}}} g_1)^T & \cdots & (\nabla_{\overline{\boldsymbol{H}}} g_r)^T \\ -\lambda_1 & \lambda_1 \nabla_{\overline{\boldsymbol{H}}} g_1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ -\lambda_r & \lambda_r \nabla_{\overline{\boldsymbol{H}}} g_r & 0 & \cdots & 0 \end{vmatrix}, \; |B| = \begin{vmatrix} g_{r+1} - y & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & g_m - y \end{vmatrix}. \tag{4}$$

As a result, we have $\det(\nabla_{\boldsymbol{w}}\boldsymbol{F}(t,\boldsymbol{w}^*)) = |A| \cdot |B|$. Obviously, $|B| \neq 0$. Furthermore, we decompose $|A|$ by $r \geq 1$ and elementary row operations to obtain

$$\begin{aligned} |A| &= \begin{vmatrix} 0 & \mathbf{0} & -1 & \cdots & -1 \\ \mathbf{0} & \sum_{i=1}^{m} \lambda_i \nabla^2_{\overline{\boldsymbol{H}}} g_i & (\nabla_{\overline{\boldsymbol{H}}} g_1)^T & \cdots & (\nabla_{\overline{\boldsymbol{H}}} g_r)^T \\ -\lambda_1 & \lambda_1 \nabla_{\overline{\boldsymbol{H}}} g_1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ -\sum_{i=1}^{r} \lambda_i & \sum_{i=1}^{r} \lambda_i \nabla_{\overline{\boldsymbol{H}}} g_i & 0 & \cdots & 0 \end{vmatrix} = \begin{vmatrix} 0 & \mathbf{0} & -1 & \cdots & -1 \\ \mathbf{0} & \sum_{i=1}^{m} \lambda_i \nabla^2_{\overline{\boldsymbol{H}}} g_i & (\nabla_{\overline{\boldsymbol{H}}} g_1)^T & \cdots & (\nabla_{\overline{\boldsymbol{H}}} g_r)^T \\ -\lambda_1 & \lambda_1 \nabla_{\overline{\boldsymbol{H}}} g_1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ -1 & 0 & 0 & \cdots & 0 \end{vmatrix} \\ &= (-1)^{N+r+1} \begin{vmatrix} \mathbf{0} & -1 & \cdots & -1 \\ \sum_{i=1}^{m} \lambda_i \nabla^2_{\overline{\boldsymbol{H}}} g_i & (\nabla_{\overline{\boldsymbol{H}}} g_1)^T & \cdots & (\nabla_{\overline{\boldsymbol{H}}} g_r)^T \\ \lambda_1 \nabla_{\overline{\boldsymbol{H}}} g_1 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ \lambda_{r-1} \nabla_{\overline{\boldsymbol{H}}} g_{r-1} & 0 & \cdots & 0 \end{vmatrix}. \end{aligned}$$

Since $\lambda_r \neq 0$, by elementary column operations, we have

$$\begin{aligned} |A| &= \frac{(-1)^{N+r+1}}{\lambda_r} \begin{vmatrix} \mathbf{0} & -1 & \cdots & -\sum_{i=1}^{r} \lambda_r \\ \sum_{i=1}^{m} \lambda_i \nabla^2_{\overline{\boldsymbol{H}}} g_i & (\nabla_{\overline{\boldsymbol{H}}} g_1)^T & \cdots & \sum_{i=1}^{r} \lambda_r (\nabla_{\overline{\boldsymbol{H}}} g_r)^T \\ \lambda_1 \nabla_{\overline{\boldsymbol{H}}} g_1 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ \lambda_{r-1} \nabla_{\overline{\boldsymbol{H}}} g_{r-1} & 0 & \cdots & 0 \end{vmatrix} \\ &= -\frac{1}{\lambda_r} \begin{vmatrix} \sum_{i=1}^{m} \lambda_i \nabla^2_{\overline{\boldsymbol{H}}} g_i & (\nabla_{\overline{\boldsymbol{H}}} g_1)^T & \cdots & (\nabla_{\overline{\boldsymbol{H}}} g_{r-1})^T \\ \lambda_1 \nabla_{\overline{\boldsymbol{H}}} g_1 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ \lambda_{r-1} \nabla_{\overline{\boldsymbol{H}}} g_{r-1} & 0 & \cdots & 0 \end{vmatrix} \\ &= -\frac{1}{\lambda_r} \prod_{i=1}^{r-1} \lambda_i \begin{vmatrix} \sum_{i=1}^{m} \lambda_i \nabla^2_{\overline{\boldsymbol{H}}} g_i & (\nabla_{\overline{\boldsymbol{H}}} g_1)^T & \cdots & (\nabla_{\overline{\boldsymbol{H}}} g_{r-1})^T \\ \nabla_{\overline{\boldsymbol{H}}} g_1 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ \nabla_{\overline{\boldsymbol{H}}} g_{r-1} & 0 & \cdots & 0 \end{vmatrix} = -\frac{1}{\lambda_r} \prod_{i=1}^{r-1} \lambda_i \cdot \begin{vmatrix} C_1 & C_2^T \\ C_2 & \mathbf{0} \end{vmatrix}, \end{aligned}$$

where $C_1 = \sum\limits_{i=1}^{m} \lambda_i \nabla^2_{\overline{\boldsymbol{H}}} g_i$, and $C_2^T = ((\nabla_{\overline{\boldsymbol{H}}} g_1)^T, \cdots, (\nabla_{\overline{\boldsymbol{H}}} g_{r-1})^T)$. It follows by the condition that C_1 is invertible, and $|A|$ becomes

$$\begin{aligned}
|A| &= -\frac{1}{\lambda_r} \prod_{i=1}^{r-1} \lambda_i \cdot \begin{vmatrix} I_N & \mathbf{0} \\ C_2 C_1^{-1} & I_{r-1} \end{vmatrix} \cdot \begin{vmatrix} C_1 & \mathbf{0} \\ \mathbf{0} & -C_2 C_1^{-1} C_2^T \end{vmatrix} \cdot \begin{vmatrix} I_N & C_1^{-1} C_2^T \\ \boldsymbol{O} & I_{r-1} \end{vmatrix} \\
&= -\frac{1}{\lambda_r} \prod_{i=1}^{r-1} \lambda_i \cdot \begin{vmatrix} C_1 & \mathbf{0} \\ \mathbf{0} & -C_2 C_1^{-1} C_2^T \end{vmatrix} = -\frac{1}{\lambda_r} \prod_{i=1}^{r-1} \lambda_i \cdot |C_1| \cdot \left| -C_2 C_1^{-1} C_2^T \right|,
\end{aligned}$$

where I_N and I_{r-1} are N-th and $(r-1)$-th identity matrices. Hence, the determinant $\det(\nabla_{\boldsymbol{w}} \boldsymbol{F}(t, \boldsymbol{w}^*))$ is simplified as

$$\det(\nabla_{\boldsymbol{w}} \boldsymbol{F}(t, \boldsymbol{w}^*)) = -\frac{1}{\lambda_r} \prod_{i=1}^{r-1} \lambda_i \cdot |C_1| \cdot \left| -C_2 C_1^{-1} C_2^T \right| \cdot |B|. \tag{5}$$

(ii) Next, we show that the rank of $C_2 C_1^{-1} C_2^T$ is $r-1$.
By $LICQ$, the number of active constraints must be less than or equals to the number of variables, that is, $r \leq N+1$, which implies $r-1 \leq N$. Hence, if $\operatorname{rank}(C_2) = r-1$, we can have $\operatorname{rank}(C_2 C_1^{-1} C_2^T) = r-1$. For this purpose, we prove that $\{\nabla_{\overline{\boldsymbol{H}}} g_1, \nabla_{\overline{\boldsymbol{H}}} g_2, \cdots, \nabla_{\overline{\boldsymbol{H}}} g_{r-1}\}$ are linearly independent. Since the gradients $\{\nabla_{(y, \overline{\boldsymbol{H}})}(g_i(\overline{\boldsymbol{H}}_t^*) - y_t^*), i = 1, 2, \cdots, r\}$ are linearly independent, the rank of the matrix

$$D = \begin{pmatrix} -1 & -1 & \cdots & -1 \\ (\nabla_{\overline{\boldsymbol{H}}} g_1)^T & (\nabla_{\overline{\boldsymbol{H}}} g_2)^T & \cdots & (\nabla_{\overline{\boldsymbol{H}}} g_r)^T \end{pmatrix}$$

is r. By a series of elementary column operations, we obtain

$$\begin{aligned}
& D \begin{pmatrix} 1 & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \\ 0 & \cdots & 0 & \lambda_r \end{pmatrix} \begin{pmatrix} 1 & 0 & \cdots & \lambda_1 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & \lambda_2 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \cdots \begin{pmatrix} 1 & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 1 & \lambda_{r-1} \\ 0 & \cdots & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} -1 & \cdots & -1 & -1 \\ (\nabla_{\overline{\boldsymbol{H}}} g_1)^T & \cdots & (\nabla_{\overline{\boldsymbol{H}}} g_{r-1})^T & 0 \end{pmatrix}.
\end{aligned} \tag{6}$$

Similarly, by a series of elementary column operations, we obtain

$$\begin{aligned}
& \begin{pmatrix} -1 & \cdots & -1 & -1 \\ (\nabla_{\overline{\boldsymbol{H}}} g_1)^T & \cdots & (\nabla_{\overline{\boldsymbol{H}}} g_{r-1})^T & 0 \end{pmatrix} \begin{pmatrix} 1 & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \\ -1 & \cdots & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & -1 & \cdots & 1 \end{pmatrix} \cdots \begin{pmatrix} 1 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & -1 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 0 & \cdots & 0 & -1 \\ (\nabla_{\overline{\boldsymbol{H}}} g_1)^T & \cdots & (\nabla_{\overline{\boldsymbol{H}}} g_{r-1})^T & 0 \end{pmatrix}.
\end{aligned} \tag{7}$$

It follows by (6) and (7) that the rank of $\left((\nabla_{\overline{\boldsymbol{H}}} g_1)^T, (\nabla_{\overline{\boldsymbol{H}}} g_2)^T, \cdots, (\nabla_{\overline{\boldsymbol{H}}} g_{r-1})^T\right)$ is $r-1$, which is full rank. Hence, the rank of C_1 is $N \geq r-1$, and then the rank of the matrix $C_2 C_1^{-1} C_2^T$ is also $r-1$, which is full rank. Thus,

$$|C_2 C_1^{-1} C_2^T| \neq 0. \tag{8}$$

(iii) Consequently, by (4), (5) and (8), we have $|\nabla_{\boldsymbol{w}} \boldsymbol{F}(t, \boldsymbol{w}^*)| \neq 0$. This completes the proof. □

Based on the discussion above, the optimality condition is a group of equations in the set Γ for each t. Hence, the optimal $\boldsymbol{w}$ is a function of t. To discuss whether this function is continuous, we have the following theorem.

Theorem 3.3 *Assume that the conditions in Theorem 3.2 hold for any t in Ω, and $\boldsymbol{F}$ and $\nabla_{\boldsymbol{w}}\boldsymbol{F}$ are continuous in $U \times V \subset \mathbb{R} \times \mathbb{R}^{N+m+1}$, which is a neighborhood of the point $(t, \boldsymbol{w})$. Then, there exists a neighborhood $U_0 \times V_0$ of $(t^*, \boldsymbol{w}^*)$ in $U \times V$, and a unique continuous function $\boldsymbol{\psi} : U_0 \to V_0$ such that $\boldsymbol{F}(t, \boldsymbol{\psi}(t)) = \boldsymbol{0},\ t \in U_0$.*

Proof First, we denote the closed interval with center t^* and distance d by $\overline{B}(t^*, d)$, and the closed ball with center $\boldsymbol{w}^*$ and distance δ by $\overline{B}(\boldsymbol{w}^*, \delta)$. Then, the set of continuous vector functions which are defined in $\overline{B}(t^*, d)$ and take value in $\overline{B}(\boldsymbol{w}^*, \delta)$ is denoted by $C(\overline{B}(t^*, d), \overline{B}(\boldsymbol{w}^*, \delta))$.

By Theorem 3.1, the matrix $\nabla_{\boldsymbol{w}}\boldsymbol{F}(t^*, \boldsymbol{w}^*)$ is invertible. Hence, we have

$$\left(\nabla_{\boldsymbol{w}}\boldsymbol{F}(t^*, \boldsymbol{w}^*)\right)^{-1} \nabla_{\boldsymbol{w}}\boldsymbol{F}(t^*, \boldsymbol{w}^*) = I. \tag{9}$$

Define a mapping $T : \boldsymbol{\psi} \to T\boldsymbol{\psi}$ as:

$$(T\boldsymbol{\psi})(t) := \boldsymbol{\psi}(t) - \left(\nabla_{\boldsymbol{w}}\boldsymbol{F}(t^*, \boldsymbol{w}^*)\right)^{-1} \boldsymbol{F}(t, \boldsymbol{\psi}(t)), t \in \overline{B}(t^*, d).$$

For any $\boldsymbol{\psi}(t), \boldsymbol{\phi}(t) \in C(\overline{B}(t^*, d), \overline{B}(\boldsymbol{w}^*, \delta))$, we define its norm as

$$\|\boldsymbol{\psi}(t) - \boldsymbol{\phi}(t)\|_\infty := \max_{q \in \{1,2,\cdots,N+m+1\}} \max_{t \in \overline{B}(t^*,d)} |\psi_q(t) - \phi_q(t)|.$$

Since $\nabla_{\boldsymbol{w}}\boldsymbol{F}$ is continuous, it follows by (9) that $\forall \epsilon > 0$, there exist $d > 0, \delta > 0$, and $\boldsymbol{\varphi}^1, \cdots, \boldsymbol{\varphi}^{N+m+1} \in C(\overline{B}(t^*, d), \overline{B}(\boldsymbol{w}^*, \delta))$, such that

$$\left\| I - \left(\nabla_{\boldsymbol{w}}\boldsymbol{F}(t^*, \boldsymbol{w}^*)\right)^{-1} \boldsymbol{G}(\boldsymbol{\varphi}^1, \cdots, \boldsymbol{\varphi}^{N+m+1}) \right\|_\infty < \epsilon,$$

where $\|.\|_\infty$ is the maximum norm of matrix, and

$$\boldsymbol{G}(\boldsymbol{\varphi}^1, \cdots, \boldsymbol{\varphi}^{N+m+1}) := \begin{pmatrix} \nabla_{\boldsymbol{w}} F_1(t, \boldsymbol{\varphi}^1) \\ \vdots \\ \nabla_{\boldsymbol{w}} F_{N+m+1}(t, \boldsymbol{\varphi}^{N+m+1}) \end{pmatrix}.$$

Hence, for any two vector functions $\boldsymbol{\psi}(t), \boldsymbol{\phi}(t)$ in $C(\overline{B}(t^*, d), \overline{B}(\boldsymbol{w}, \delta))$, we have

$$T\boldsymbol{\psi} - T\boldsymbol{\phi} = \boldsymbol{\psi}(t) - \boldsymbol{\phi}(t) - \left[\left(\nabla_{\boldsymbol{w}}\boldsymbol{F}(t^*, \boldsymbol{w}^*)\right)^{-1} \left(\boldsymbol{F}(t, \boldsymbol{\psi}(t)) - \boldsymbol{F}(t, \boldsymbol{\phi}(t))\right) \right].$$

Based on the mean value theorem, for any $q \in \{1, 2, \ldots, N + m + 1\}$, there exists $\eta_q \in (0, 1)$, such that

$$F_q(t, \boldsymbol{\psi}(t)) - F_q(t, \boldsymbol{\phi}(t)) = \nabla_{\boldsymbol{w}} F_q(t, \boldsymbol{\varphi}^q(t)(\boldsymbol{\psi}(t) - \boldsymbol{\phi}(t)),$$

where $\boldsymbol{\varphi}^q(t) = \eta_q \boldsymbol{\psi}(t) + (1 - \eta_q)\boldsymbol{\phi}(t)$. Note that

$$\begin{aligned} &\|\boldsymbol{\varphi}^q(t) - \boldsymbol{w}^*\|_\infty = \|\eta_q(\boldsymbol{\psi}(t) - \boldsymbol{w}^*) + (1 - \eta_q)(\boldsymbol{\phi}(t) - \boldsymbol{w}^*)\|_\infty \\ \le &\eta_q \|\boldsymbol{\psi}(t) - \boldsymbol{w}^*\|_\infty + (1 - \eta_q)\|\boldsymbol{\phi}(t) - \boldsymbol{w}^*)\|_\infty \le \eta_q \delta + (1 - \eta_q)\delta = \delta, \end{aligned}$$

we obtain $\boldsymbol{\varphi}^q(t) \in C(\overline{B}(t^*, d), \overline{B}(\boldsymbol{w}^*, \delta))$, $\forall q = 1, 2, \cdots, N+m+1$. Then,

$$\begin{aligned}
&(\nabla_{\boldsymbol{w}} \boldsymbol{F}(t^*, \boldsymbol{w}^*))^{-1} (\boldsymbol{F}(t, \boldsymbol{\psi}(t)) - \boldsymbol{F}(t, \boldsymbol{\phi}(t))) \\
=&(\nabla_{\boldsymbol{w}} \boldsymbol{F}(t^*, \boldsymbol{w}^*))^{-1} \begin{bmatrix} \nabla_{\boldsymbol{w}} F_1(t, \boldsymbol{\varphi}^1) \\ \vdots \\ \nabla_{\boldsymbol{w}} F_{N+m+1}(t, \boldsymbol{\varphi}^{N+m+1}) \end{bmatrix} (\boldsymbol{\psi}(t) - \boldsymbol{\phi}(t)) \\
=&(\nabla_{\boldsymbol{w}} \boldsymbol{F}(t^*, \boldsymbol{w}^*))^{-1} \boldsymbol{G}(\varphi^1, \cdots, \boldsymbol{\varphi}^{N+m+1})(\boldsymbol{\psi}(t) - \boldsymbol{\phi}(t)).
\end{aligned}$$

Hence, we have

$$\begin{aligned}
&\|T\boldsymbol{\psi} - T\boldsymbol{\phi}\|_\infty \\
=&\left\| \left(I - (\nabla_{\boldsymbol{w}} \boldsymbol{F}(t^*, \boldsymbol{w}^*))^{-1} \boldsymbol{G}(\boldsymbol{\varphi}^1, \cdots, \boldsymbol{\varphi}^{N+m+1}) \right) (\boldsymbol{\psi}(t) - \boldsymbol{\phi}(t)) \right\|_\infty \\
\le&(N+m+1) \left\| I - (\nabla_{\boldsymbol{w}} \boldsymbol{F}(t^*, \boldsymbol{w}^*))^{-1} \boldsymbol{G}(\boldsymbol{\varphi}^1, \cdots, \boldsymbol{\varphi}^{N+m+1}) \right\|_\infty \|(\boldsymbol{\psi}(t) - \boldsymbol{\phi}(t))\|_\infty \\
\le&(N+m+1) \cdot \epsilon \cdot \|(\boldsymbol{\psi}(t) - \boldsymbol{\phi}(t))\|_\infty.
\end{aligned}$$

Set $\epsilon < \frac{1}{2(N+m+1)}$, it implies that

$$\|T\boldsymbol{\psi} - T\boldsymbol{\phi}\|_\infty < \frac{1}{2} \|\boldsymbol{\psi} - \boldsymbol{\phi}\|_\infty.$$

Therefore, T is a compression mapping in the space $C(\overline{B}(t^*, d), \overline{B}(\boldsymbol{w}^*, \delta))$.

Furthermore, it follows from (9) and the continuity of $\boldsymbol{F}(t, \boldsymbol{w})$ that there exists $d_1 > 0$, such that if $0 < d < d_1$, then

$$\begin{aligned}
&\left\| (\nabla_{\boldsymbol{w}} \boldsymbol{F}(t^*, \boldsymbol{w}^*))^{-1} \boldsymbol{F}(t, \boldsymbol{w}^*) \right\|_\infty \\
=&\left\| (\nabla_{\boldsymbol{w}} \boldsymbol{F}(t^*, \boldsymbol{w}^*))^{-1} \boldsymbol{F}(t, \boldsymbol{w}^*) - \boldsymbol{F}(t^*, \boldsymbol{w}^*) \right\|_\infty < \frac{\delta}{2}, \quad \forall t \in \overline{B}(t^*, d). \qquad (10)
\end{aligned}$$

For the simplicity of notation, let $\boldsymbol{w}^*$ denote the constant function defined by $\overline{B}(t^*, d) = \boldsymbol{w}^*$. Then, $\boldsymbol{w}^* \in C(\overline{B}(t^*, d), \overline{B}(\boldsymbol{w}^*, \delta))$, and

$$T\boldsymbol{w}^* = \boldsymbol{w}^* - (\nabla_{\boldsymbol{w}} \boldsymbol{F}(t^*, \boldsymbol{w}^*))^{-1} \boldsymbol{F}(t, \boldsymbol{w}^*). \qquad (11)$$

By (10) and (11), we have

$$\|T\boldsymbol{w}^* - \boldsymbol{w}^*\|_\infty = \left\| (\nabla_{\boldsymbol{w}} \boldsymbol{F}(t^*, \boldsymbol{w}^*))^{-1} \boldsymbol{F}(t, \boldsymbol{w}^*) \right\|_\infty < \frac{\delta}{2}.$$

Thus, for any function $\boldsymbol{\psi} \in C(\overline{B}(t^*, d), \overline{B}(\boldsymbol{w}^*, \delta))$ which satisfies $\boldsymbol{\psi}(t^*) = \boldsymbol{w}^*$, if $0 < d < d_1$, we obtain

$$\begin{aligned}
\|T\boldsymbol{\psi} - \boldsymbol{w}^*\|_\infty &\le \|T\boldsymbol{\psi} - T\boldsymbol{w}^*\|_\infty + \|T\boldsymbol{w}^* - \boldsymbol{w}^*\|_\infty \\
&\le \frac{1}{2} \|\boldsymbol{\psi} - \boldsymbol{w}^*\|_\infty + \frac{\delta}{2} \le \frac{\delta}{2} + \frac{\delta}{2} = \delta.
\end{aligned}$$

Let $X := C(\overline{B}(t^*, d), \overline{B}(\boldsymbol{w}^*, \delta))$, then $(X, \|.\|_\infty)$ is a Banach space, which is complete. Since T is a compression mapping and $TX \subset X$, it follows from fixed-point theorem that there exists a unique $\boldsymbol{\psi} \in X$ such that $T\boldsymbol{\psi} = \boldsymbol{\psi}$, i.e.,

$$\boldsymbol{\psi}(t) - (\nabla_{\boldsymbol{w}} \boldsymbol{F}(t^*, \boldsymbol{w}^*))^{-1} F(t, \boldsymbol{\psi}(t)) = \boldsymbol{\psi}(t), \quad \forall t \in \overline{B}(t^*, d).$$

Then,

$$(\nabla_{\boldsymbol{w}} \boldsymbol{F}(t^*, \boldsymbol{w}^*))^{-1} F(t, \boldsymbol{\psi}(t)) = \mathbf{0}, \quad \forall t \in \overline{B}(t^*, d).$$

$$\boldsymbol{F}(t, \boldsymbol{\psi}(t)) = (\nabla_{\boldsymbol{w}} \boldsymbol{F}(t^*, \boldsymbol{w}^*)) \cdot \mathbf{0} = \mathbf{0}, \quad \forall t \in \overline{B}(t^*, d).$$

Note that $\boldsymbol{\psi}(t)$ is unique, we obtain $\boldsymbol{\psi}(t^*) = \boldsymbol{w}^*$. This completes the proof. □

By Theorem 3.3, the function $\boldsymbol{\psi}(t), t \in \overline{B}(t^*, d)$ is defined only in a neighborhood of t^*, which is not Ω. It can be generalized to Ω by the uniqueness of $\boldsymbol{\psi}(t)$ and finite covering theorem easily, which is stated by the following theorem.

Theorem 3.4 *Assume that the conditions in Theorem 3.3 hold for any $t \in \Omega$, then, there exists a unique continuous function $\boldsymbol{\psi}(t)$ which is defined in Ω, such that $\boldsymbol{F}(t, \boldsymbol{\psi}(t)) = 0, \forall t \in \Omega$.*

By using the above two theorems, we finally arrive at the following theorem.

Theorem 3.5 *Assume that the conditions in Theorem 3.3 hold for any t in Ω, then,*

$$y^* = \lim_{L\to\infty} y^{*L} = \max_{t\in\Omega} y_t^* = y^{*t}. \tag{12}$$

Proof By Lemma 3.1, y_t^* and $\overline{\boldsymbol{H}}^*(t)$ satisfy

$$\max_{i\in\{1,\ldots,m\}} g_i(\overline{\boldsymbol{H}}^*(t)) - y_t^* = 0, \quad \forall t \in \Omega.$$

By Theorem 3.4, the function $\overline{\boldsymbol{H}}^*(t)$ is continuous at Ω. Thus, there exists a series $\{\boldsymbol{H}^{(L)}(t) : L = 1, 2, \cdots\}$ with each $\boldsymbol{H}^{(L)}(t)$ in $\mathrm{span}^N(\Delta_L)$ such that

$$\lim_{L\to\infty} \boldsymbol{H}^{(L)}(t) = \overline{\boldsymbol{H}}^*(t), \quad \forall t \in \Omega.$$

Define

$$y^L = \max_{t\in\Omega} \max_i g_i(\boldsymbol{H}^{(L)}(t)).$$

Then, by the optimality of Problem $P(L)$, we must have $y^L \geq y^{*L}$. By the continuity of the functions $g_i, i = 1, \ldots, m$, we have

$$y^* = \lim_{L\to\infty} y^L = \lim_{L\to\infty} \max_{t\in\Omega} \max_i g_i(\boldsymbol{H}^{(L)}(t)) = \max_{t\in\Omega} \max_i g_i(\overline{\boldsymbol{H}}^*(t)) = \max_{t\in\Omega} y_t^*.$$

Note that $y^L \geq y^{*L}$, we have

$$\lim_{L\to\infty} y^{*L} \leq \lim_{L\to\infty} y^L = \max_{t\in\Omega} y_t^*. \tag{13}$$

On the other hand, since $\overline{\boldsymbol{H}}^{*L}(t) \in \text{span}^N(\Delta_L)$ is continuous, and $\overline{\boldsymbol{H}}^{*}(t)$ is continuous and optimal, we have $y^{*L} \geq \max_{t\in\Omega} y_t^*$. Take the limit and we obtain

$$\lim_{L\to\infty} y^{*L} \geq \max_{t\in\Omega} y_t^*. \tag{14}$$

Thus, by (13) and (14), the equation (12) holds. This completes the proof. □

By Theorem 3.5, we can compute the limit of the optimal values by solving Problem $P(t)$ for every t. Note that the parameter L is not present and therefore the computational complexity can be reduced, compared with the general method for choosing a sufficiently large L.

4 Numerical Experiments

In this section, we verify Theorem 3.5 by the following two examples, where the computations were implemented in Matlab.

Example 4.1*:* We consider Problem $P(L)$, where the constraints are

$$\left|\sum_{k=1}^{N} A_{ik}(t,r_i)H_k(\boldsymbol{x},t) - G_d(t)\right|^2 + \sum_{k=1}^{N} H_k^2(\boldsymbol{x},t) - y \leq 0,\ \forall t\in\Omega, i=1,\cdots,21,$$

where $\Omega = \Omega_1 \cup \Omega_2$, $G_d(t)$ equals to 1 when $t\in\Omega_1$ and 0 when $t\in\Omega_2$, and

$$A_{ik}(t,r_i) = \cos((2k-1)t + r_i), H_k(\boldsymbol{x},t) = \sum_{j=1}^{L} x_{kj}t^{j-1},\quad k=1,2,\cdots,N.$$

Note that the power function series are used as the basis functions. We set $N = 4$, $r_i = 1 + 0.05 * (i-1)$, $\Omega_1 = [0.1\pi, 0.3\pi]$, $\Omega_2 = [0.5\pi, \pi]$, and choose the parameter L from 1 to 10. By choosing a sufficiently dense grid as $0.1\pi : 0.02 : 0.3\pi$ and $0.5\pi : 0.02 : \pi$, we solve the problem and the optimal values are depicted in Figure 1. It can be seen that the optimal value series is monotonically decreasing and the limit exists.

Moreover, it is not difficult to verify that the conditions in Theorem 3.2 hold. Then, we apply the proposed method to find the limit. By choosing the same grid as above, we solve Problem $P(t)$ for each t and obtain $y^* = 0.4449$. The corresponding optimal solution is $\boldsymbol{H}^* = [-0.1135, -0.3826, -0.2815, 0.0919]^T$. Obviously, $\left|\nabla^2_{\boldsymbol{H}}\left(\left|\boldsymbol{A}_m^T(t,r_m)\boldsymbol{H}(\boldsymbol{x},t) - G_d(t)\right|^2 + \sum_{k=1}^{N} H_k^2(\boldsymbol{x},t)\right)\right| > 0$ at point $\boldsymbol{H}^*$.

It can be seen from Figure 1 that the optimal value series approaches y^*, and the relation (12) has been verified. Moreover, the running time of the proposed method is only 2.90 seconds, while it costs 377.44 seconds in discretization method when $L = 6$. Thus, the proposed is very efficient.

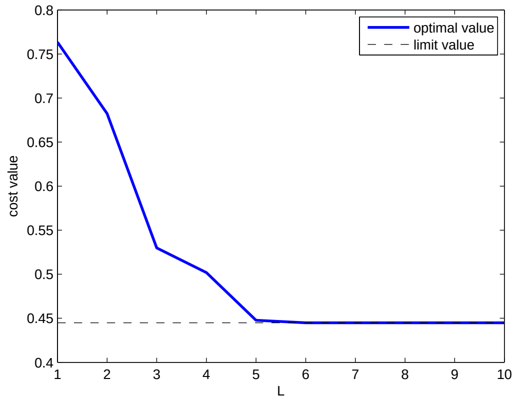

Fig. 1 The optimal values with different parameter L.

Example 4.2*:* We consider a far field microphone array beamformer design problem in [9]. The constraints of Problem $P(L)$ are

$$\left(\sum_{k=1}^{N} A_k^r(\theta,f)H_k^r(\boldsymbol{x},f) + \sum_{k=1}^{N} A_k^i(\theta,f)H_k^i(\boldsymbol{x},f) - G_0^r(\theta,f)\right)^2 +$$
$$\left(\sum_{k=1}^{N} A_k^r(\theta,f)H_k^i(\boldsymbol{x},f) + \sum_{k=1}^{N} A_k^i(\theta,f)H_k^r(\boldsymbol{x},f) - G_0^i(\theta,f)\right)^2 - y \le 0,\ \forall(\theta,f)\in\Omega,$$

where $N=9$, $r_i=-0.4+(i-1)/10, i=1,\cdots,N$, $f_s=8kHz$, $c=340.9m/s$, $\tau=(L-1)/2$, $L=1+2p$, $p=0,\cdots,20$, $\Omega_s=\Omega_{s_1}\cup\Omega_{s_2}$, $\Omega=\Omega_p\cup\Omega_s$, $(G_0^r(\theta,f),G_0^i(\theta,f))$ equals to $(\cos(2\pi f\tau/f_s),-\sin(2\pi f\tau/f_s))$ when $(\theta,f)\in\Omega_p$ and zero vector when $(\theta,f)\in\Omega_s$. For each $k=1,2,\cdots,N$,

$$A_k^r(\theta,f)=\cos(2\pi f r_k\cos\theta/c), A_k^i(\theta,f)=-\sin(2\pi f r_k\cos\theta/c),$$
$$H_k^r(\boldsymbol{x},f)=\boldsymbol{x}_k^T\boldsymbol{d}_0^r(f), H_k^i(\boldsymbol{x},f)=\boldsymbol{x}_k^T\boldsymbol{d}_0^i(f), \boldsymbol{x}_k=[x_k(0),x_k(1),\cdots,x_k(L-1)]^T,$$
$$\boldsymbol{d}_0^r(f)=[1,\cos(2\pi f/f_s),\cdots,\cos(2\pi f(L-1)/f_s)]^T,$$
$$\boldsymbol{d}_0^i(f)=[0,-\sin(2\pi f/f_s),\cdots,-\sin(2\pi f(L-1)/f_s)]^T.$$

$$\begin{aligned}\Omega_p &= \{(f,\theta): f=[0.7kHz,2.0kHz],\theta\in\{85^\circ:2.5^\circ:95^\circ\}\},\\ \Omega_{s_1} &= \{(f,\theta): f=[0.7kHz,2.0kHz],\theta\in\{0^\circ:2.5^\circ:50^\circ\}\cup\{130^\circ:2.5^\circ:180^\circ\}\},\\ \Omega_{s_2} &= \{(f,\theta):\theta\in\{0^\circ:2.5^\circ:50^\circ\}\cup\{85^\circ:2.5^\circ:95^\circ\}\cup\{130^\circ:2.5^\circ:180^\circ\},\\ &\qquad f=[2.5kHz,4.0kHz]\}.\end{aligned}$$

First of all, we can verify that the conditions in Theorem 3.2 hold. Then, by choosing a sufficiently dense grid $0.7:0.02:2$ and $2.5:0.02:4$, we apply the proposed method to obtain the maximum as $y^*=3.40\times10^{-4}$. The parameter L is then varied from 1 to 41 with increment by 2, we solve for the optimal values and y^* are depicted in Figure 2. It can be seen that the optimal value series is monotonically decreasing and approaches to y^*. Hence, the relation (12) is true. Furthermore, the running time of the discretization method is 1229.70 seconds when $L = 29$, while it only costs 4.66 seconds by the proposed method. Thus, the proposed method is very efficient.

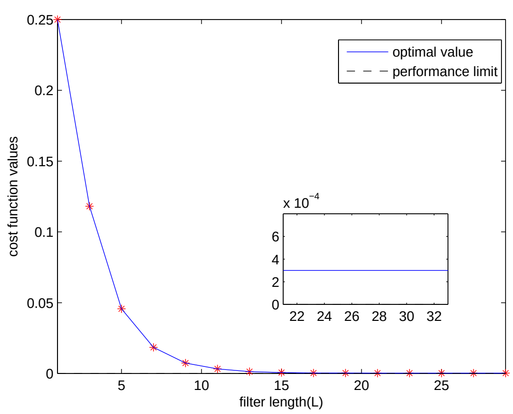

Fig. 2 The optimal values with different parameter L.

5 Conclusions

In this paper, we have considered a class of SIP problems with a parameter L. We have analyzed the limit of the optimal value series as the parameter L tends to infinity. To compute the limit efficiently, we proposed a novel method which decomposes the problem into a series of simplified subproblems. By taking the maximum of these optimal values, the final limit can be obtained. Furthermore, we have derived the conditions and have applied a fixed-point theorem to support the theoretical basis of the proposed method. Finally, we have verified the efficiency of the proposed method by numerical examples.

Acknowledgements This paper is supported by the grant of National Natural Science Foundation of China (No. 11771064,11431004), the grant of Guangdong Basic and Applied Basic Research Foundation (No. 2020A1515010463), the program for scientific research start-up funds of Guangdong Ocean University, the grant of Chongqing Normal University (No. 17XLB010). The fourth author is supported by RGC Grant PolyU. (152245/18E) and PolyU grant ZZGS.

References

1. Hettich, R., Kortanek, K.O.: Semi-infinite programming: Theory, methods, and application. SIAM Review **35**, 380–429 (1993)

2. López, M., Still, G.: Semi-infinite programming. European Journal of Operational Research **180**, 491–518 (2007)
3. Shapiro, A.: Semi-infinite programming, duality, discretization and optimality conditions. Optimization **58**, 133–161 (2009)
4. Feng, Z.G., Teo, K.L., Rehbock, V.: A smoothing approach for semi-infinite programming with projected new-type algorithm. Journal of Industrial and Management Optimization **5**, 141–151 (2009)
5. Shiu, T.J., Wu, S.Y.: Relaxed cutting plane method with convexification for solving nonlinear semi-infinite programming. Computational Optimization and Applications **53**(1), 91–113 (2012)
6. Wang, S.X., Yuan, Y.X.: Feasible method for convex semi-infinite programming. SIAM Journal on Optimization **25**, 2537–2560 (2010)
7. Zhang, L.P., Wu, S.Y., Lopez, M.A.: A new exchange method for convex semi-infinite programming. SIAM Journal on Optimization **20**, 2959–2977 (2010)
8. Yiu, K.F.C., Yang, X.Q., Nordholm, S., Teo, K.L.: Near-field broadband beamformer design via multidimensional semi-infinite linear programming techniques. IEEE Transactions on Speech and Audio Processing **11**(6), 725–732 (2003)
9. Feng, Z.G., Yiu, K.F.C., Nordholm, S.: Performance limit of broadband beamformer designs in space and frequency. Journal of Optimization Theory and Application **164**, 316–341 (2015)